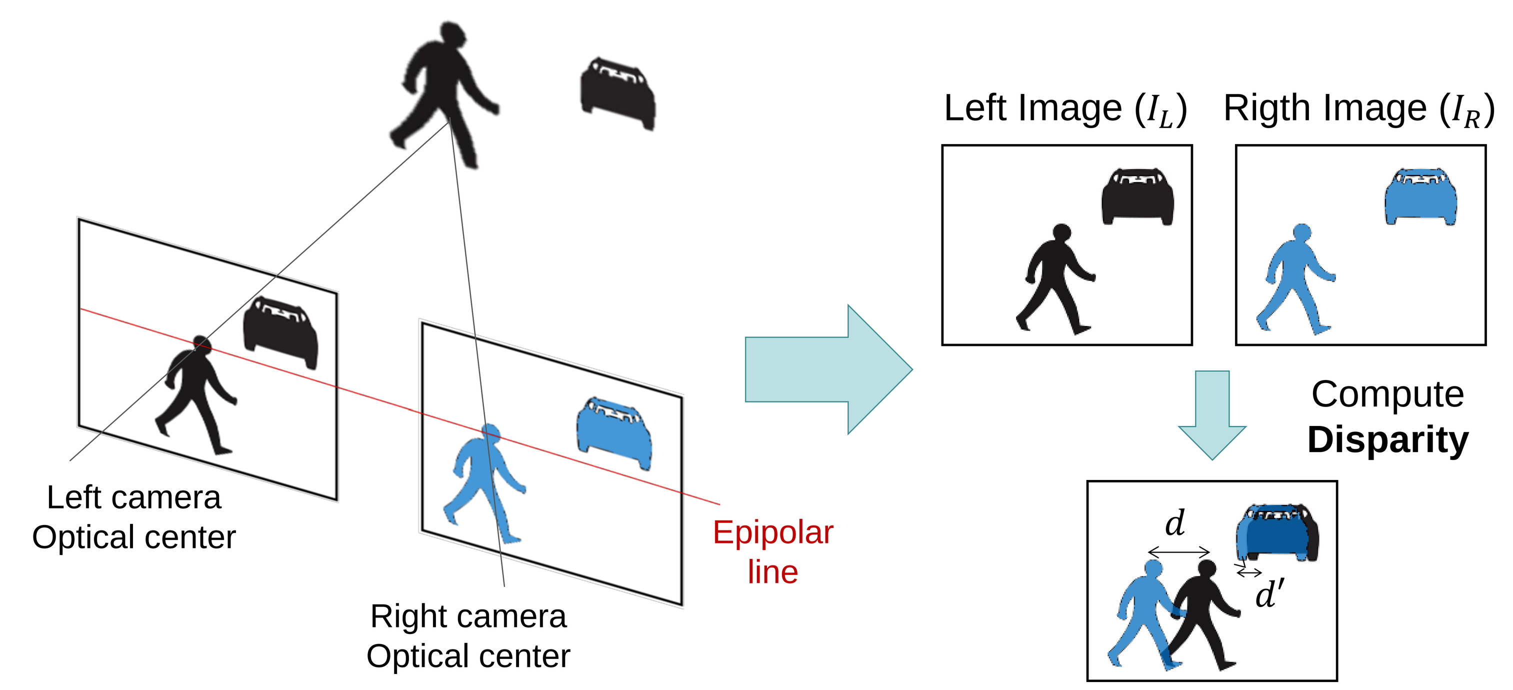
Real-time 3D Reconstruction for Autonomous Driving via Semi-Global Matching

D. Hernandez-Juarez[1], A. Chacon[1], A. Espinosa[1], D. Vazquez[2], J. C. Moure[1], A. M. Lopez[2]
Universitat Autònoma de Barcelona
[1]Computer Architecture & Operating Systems Dpt. — [2]Computer Vision Center

Abstract

Robust and dense computation of depth information from stereo-camera systems is a computationally demanding requirement for real-time autonomous driving. Semi-Global Matching (SGM) [1] approximates heavy-computation global algorithms results but with lower computational complexity, therefore it is a good candidate for a real-time implementation. SGM minimizes energy along several 1D paths across the image. The aim of this work is to provide a real-time system producing reliable results on energy-efficient hardware. Our design runs on a *NVIDIA Titan X* GPU at 104.62 FPS and on a *NVIDIA Drive PX* at 6.7 FPS, promising for real-time platforms.

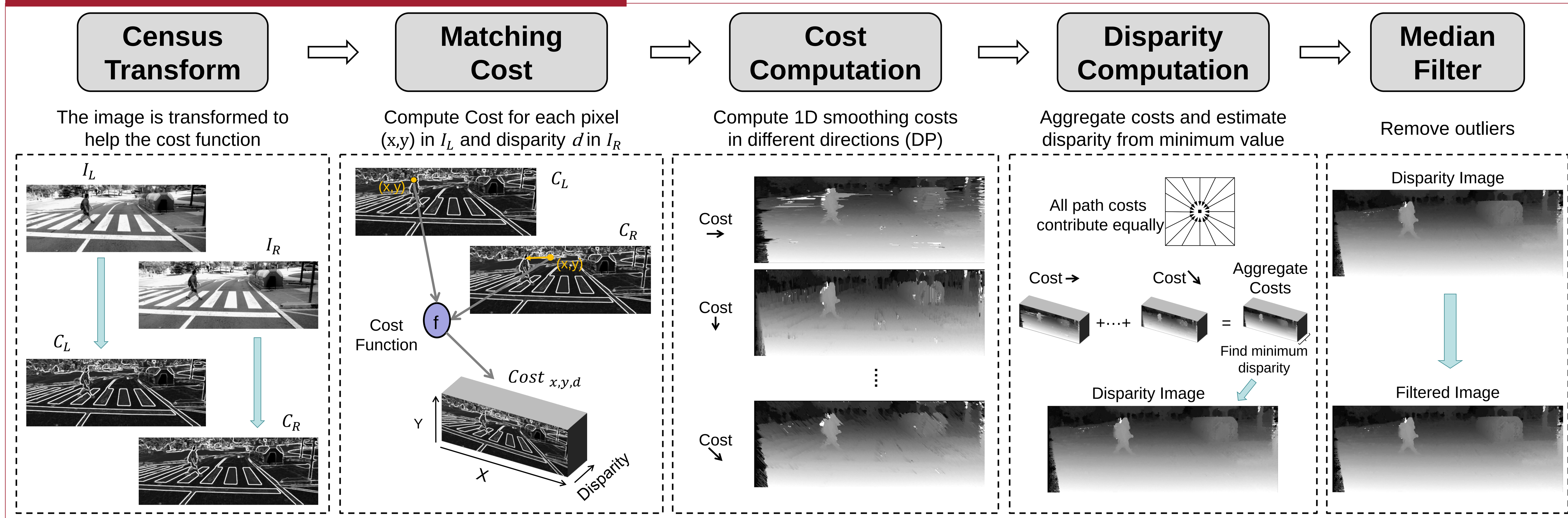
Problem: 3D Reconstruction



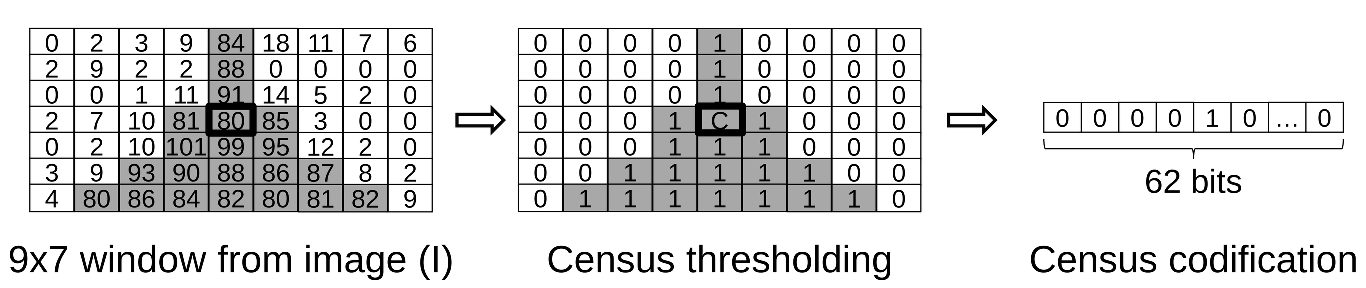
Infer **depth** from left & right images using triangulation

Disparity, d: point (x, y) in I_L corresponds to point (x-d, y) in I_R

Pipeline



Census Transform



$$Census\,(\,I\,(x,y)\,) = \begin{cases} 1, if\ I(x,y) \geq central\ value \\ 0, if\ I(x,y) < central\ value \end{cases}$$

Matching Cost: Hamming Distance

$$Cost_{x,y,d} = \text{Hamming Distance}(\,Census(\,I_L(x,y)\,), Census(\,I_R(x-d,y))\,)$$

\# of different bits

Hamming(| 1 | 0 | 1 | 0 | 0 | 1 | ... | 0 | , | 1 | 0 | 0 | 0 | 0 | 0 | ... | 0 |) = 2

Implementation: Hamming(a, b) = bitcount (a *xor* b)

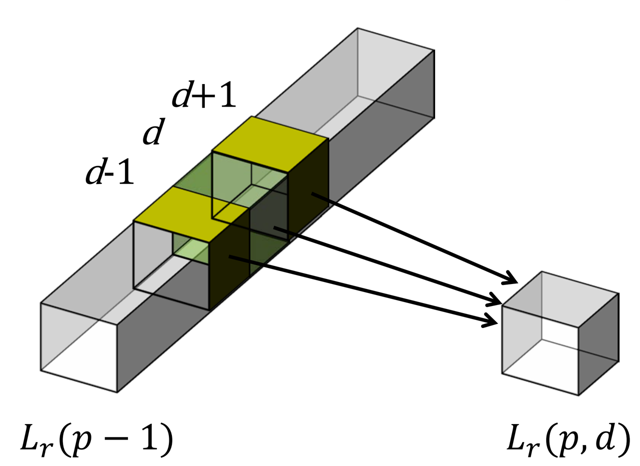
Cost Computation

1D Smoothing function: Dynamic Programming

$$L_r(p,d) = Cost(p,d) + \min \begin{cases} L_r(p-1,d) \\ L_r(p-1,d+1) + P_1 \\ L_r(p-1,d-1) + P_1 \\ min_i L_r(p-1,i) + P_2 \end{cases}$$

L_r: smoothed cost in direction r, p: pixel position, d: disparity
P_1, P_2: penalties for small and high disparity changes

GPU Implementation Details



Stencil pattern in 8 directions
1 kernel per direction
1 warp per plane
Use of registers storage
Shuffle for intra-warp communication

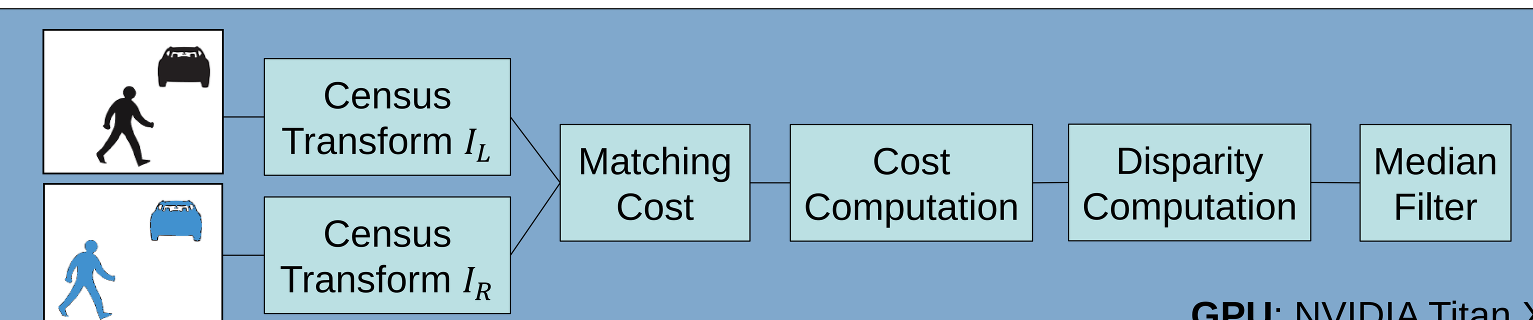
Results

	FPS	Speed Up	FPS / Watt
CPU[1] SIMD	2.3	1	0.02
GPU Naive	25.4	10.98	0.10
GPU Optimized	**104.62**	**45.49**	**0.42**
NVIDIA Drive PX[2]	**6.7**	**2.90**	**0.67**

CPU: Intel Core i7-5930K
GPU: NVIDIA Titan X
Drive PX: NVIDIA Tegra X1

Image Size: 1280x480
Disparity: 128

[1] *single-thread* [2] *single-socket*



GPU: NVIDIA Titan X

% Time	**3.15 %**	**3.67 %**	**68.38 %**	**22.60 %**	**2.20 %**
% Instructions	4.37 %	3.34 %	68.62 %	21.77 %	1.90 %
Bandwidth	19 GB/s	234 GB/s	177 GB/s	288 GB/s	12 GB/s

Conclusions:
- Semi-global matching can be used for real-time 3D reconstruction.
- Need new strategies to get real-time performance for NVIDIA Drive PX.
- NVIDIA Drive PX has 1.57x better energetic efficiency than high-end GPUs.

References:
[1] H. Hirschmüller, “Stereo Processing by Semiglobal Matching and Mutual Information,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 328–341, 2008.

Acknowledgements:
This work is supported by the Spanish MICINN projects TRA2014-57088-C2-1-R, TIN2014-53234-C2-1-R and by the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya (2014-SGR-1506) and DGT project SPIP2014-01352. Our research is also kindly supported by NVIDIA Corporation in the form of different GPU hardware.